Supplementary Material For:

Reports

# AG-NGS: A powerful and user-friendly computing application for the semi-automated preparation of next-generation sequencing libraries using open liquid handling platforms

Sergio Callejas*, Rebeca Álvarez*, Alberto Benguria, and Ana Dopazo
*Spanish National Center for Cardiovascular Research (CNIC), Madrid, Spain*

*BioTechniques* 56:28-35 (January 2014) doi 10.2144/000114124
Keywords: next-generation sequencing; sample preparation; automation; liquid handling robot; software

## Instructions and guidelines for adapting AG-NGS to other protocols and liquid handling systems

The AG-NGS application can potentially be adapted to any protocol or liquid handling system. These instructions and guidelines are intended to help users adapt AG-NGS to different protocols using the Tecan EVOware software or another liquid-handling system. Adapting AG-NGS to other protocols/systems involves changes in Microsoft Excel sheets and Visual Basic for Applications (VBA) scripts, so a certain knowledge of both is required. This knowledge should be even deeper for non Tecan Evoware users wishing to adapt AG-NGS to other systems. We recommend that users consult Microsoft Excel and VBA manuals during the process or seek help from someone experienced with these programs. A good knowledge of Tecan EVOware or other specific software is necessary, since it will also be necessary to create an appropriate script in the liquid-handling software to execute the files generated by AG-NGS and to create carriers, labware, and liquid classes. It is recommended that all users watch the tutorial videos on the AG-NGS website. Although the second part of video 1 and videos 2 and 3 focus on Tecan EVOware, the process of configuration will be very similar in other systems.

## How to modify the TruSeq sample prep script for sample library preparation in AG-NGS

**For all users.** Should you wish to use a different kit than the TruSeq RNA/DNA Sample Preparation protocol by Illumina, the AG-NGS core files need to be modified. Follow these steps to access the core files.

1. Double click on the "AG-NGS" folder and then on the "CoreFiles" folder.
2. Open the TruSeq mRNA sample prep.xls file.
3. Right click on the sheet name "Input" and then select "show"; then select "Main" and "cDNA_gwl" sheets to show these sheets.

The "Input" sheet collects data entered by the user through the VBA user form (basically the number of samples, index numbers, and carrier positions). These data are transferred to cells colored in gray and yellow in the "Input" sheet. The other cells in this sheet calculate the required volumes of reagents or the number of tips boxes needed. The "Input" sheet contains all the data that will be used in the "Main" sheet.

The "Main" sheet contains all the commands that will be executed by Tecan EVOware. Each row contains a specific command, such as "aspirate", "dispense," or "transfer labware", that will be executed in that specific order. Some commands require additional parameters that can be either fixed numbers or variables acquired from the "Input" sheet.

For example, the aspirate command (A) requires the following parameters delimited by semicolons:

- RackLabel; RackID; RackType; Position; TubeID; Volume; LiquidClass;
- RackID and TubeID are optional and can be omitted. The following command (row 93 in the "Main" sheet):
- A;InputPlate;;PCR96WellPlate;5;;110;NGS_RNAseq_PCRtoRigidPlate;;;;;;;;

These parameters instruct the robot to aspirate 110 µl of a predefined liquid class called "NGS_RNAseq_PCRtoRigidPlate" from well number 5 of a plate called "InputPlate", which is a predefined "PCR96WellPlate" plate type.

In the "Main" sheet, each of the parameters included in row 93 are located in different columns (A to H). This allows us to assign a variable value to certain parameters. For instance, the value in cell G93 is the sum of cells E3 and C4 from the "Input" sheet, called "sample_vol" and "RNA_beads_sample2" respectively, plus 10 µl excess to ensure proper aspiration of the total volume. To learn more about how to use variables and relative values in Microsoft Excel, please click the help button in the application.

The most important Evoware commands included in the TruSeq RNA/DNA Sample Preparation protocols are as follows:

- Aspirates (A)
- Dispenses (D)
- Distributes a reagent (R)
- Sets DITI type (S)
- Replaces DITI (W)
- Executes a RoMa vector (B;Vector)
- Displays comment (B;"text")
- Asks user to enter a value (B;UserPrompt)
- Transfers labware from one position to another (B;Transfer_Rack)
- Starts a timer (B;StartTimer)
- Waits for a timer (B;WaitTimer)
- Initializes, starts, and stops the shaker (B;FACTS)

**For Tecan Evoware users.** Any user wishing to modify this script can delete, add, or modify rows in the "Main" sheet. Always keep in mind that commands will be executed row by row and in the same order as written in this sheet. For more information about how to use these commands, and many others, click on the help button in Tecan Evoware and then select "worklist file format" and also "Advanced worklist". Any additional variable needed can be introduced in the "Input" sheet in any empty cell.

**For non Tecan Evoware users.** If a platform that does not use Evoware software is used, then the specific commands and parameters for that platform have to be retyped in the "Main" sheet. Check the format of files to be imported in the user manual of your specific software and modify the "Main" sheet according to that specific format. Any user wishing to modify this script can also delete, add, or modify rows in the "Main" sheet. Always keep in mind that commands will be executed row by row and in the same order as written in this sheet. Any additional variable needed can be introduced in the "Input" sheet in any empty cell.

**For all users.** Regardless of the platform used, it is very important to delimit the number of samples to be processed. The "Main" sheet has been created for 96 samples; for this reason it is necessary to introduce a formula in column A, so that unused rows are flagged when the number of sample to process is less than 96. Different formulas can be used for this propose, but must always be referred in some way to the number of samples defined in the "Input" sheet. Check column A in the "Main" sheet to see some examples of these formulas. Rows flagged with the word "REMOVE" in capitals in column A in the "Main" sheet will be removed later on (when the "Send data to robot" button is selected in the user form) to generate the definitive worklist to be executed by the liquid handling system software.

Once the "Main" sheet has been modified according to the new protocol/liquid handling system, a final worklist has to be generated as explained below.

**For Tecan Evoware users.** If you are using Tecan Evoware, then this worklist has the extension .gwl. To access the .gwl file precursor sheet, click on the cDNA_gwl sheet to activate it. In this sheet, AG-NGS uses the Excel function "CONCATENATE" to merge all columns in the "Main" sheet into a single column with items delimited by semicolons. If more parameters have been introduced in the "Main" sheet, more columns can be merged in the "cDNA_gwl" sheet.

Since some commands require more parameters than others, depending on the command, a different number of columns has to be merged. Most commands accept several semicolons at the end of the sentence, so more columns than needed can be merged to compensate for those commands that need more columns, making the merging process easier. However, commands such as B;UserPrompt or B;Vector do not accept any semicolons after the last parameter; for this reason a conditional formula (containing the function "IF") has been included in cDNA_gwl sheet column A. This formula merges a different number of columns depending on the command present in each row. Modify the formula in cell A1 to merge columns from the "Main" sheet in the appropriate format (if more parameters have been introduced in the "Main" sheet, more columns can be merged in the "cDNA_gwl" sheet). For more information on how the Microsoft Excel "IF" and "CONCATENATE" functions work, please consult the help section in Microsoft Excel.

**For non Tecan Evoware users.** Access the sheet called "cDNA_gwl" by clicking on the cDNA_gwl sheet name. In this sheet, AG-NGS uses the Excel function "CONCATENATE" to merge all columns from the "Main" sheet into a single column with items delimited by semicolons (Tecan Evoware format). Most commands in Tecan Evoware accept several semicolons at the end of the sentence, so more columns than needed can be merged to compensate for those commands that need more columns, making the merging process easier. However, commands such as B;UserPrompt or B;Vector do not accept any semicolons after the last parameter; for this reason a conditional formula (containing the function "IF") has been included in cDNA_gwl sheet column A. This formula merges a different number of columns depending on the command present in each row. In a system other than Tecan Evoware, this formula must be modified according to the requirements of the system. This may involve removing the semicolons, adding new symbols and including the functionw "IF" in the formula if different commands require a different number of parameters. Check the format of the worklists in your system, and adjust the formula in cDNA_gwl cell A1 according to that format. If more parameters have been introduced in the "Main" sheet, more columns can be merged in the "cDNA_gwl" sheet. For more information on how the Microsoft Excel "IF" and "CONCATENATE" functions work, please consult the help section in Microsoft Excel.

**For all users.** Once the formula in cell A1 has been modified, click on the lower right hand corner of A1 and drag down to the last row of your script. In the current "cDNA_gwl" sheet, this is row 4619, which corresponds to row 4631 in the "Main" sheet. From the "Main" sheet, several worklists can be generated in different Microsoft Excel sheets, so that they can be executed independently. In fact, this core file in AG-NGS contains four

worklists: "cDNA_gwl", "End_Repair_gwl", "LigateAdapter_gwl", and "Enrich_DNA_Fragments_gwl". These sheets can be accessed by right clicking on the sheet names and then selecting "show".

**For Tecan Evoware users.** At this point, all changes needed to adapt to different kits have been made. Follow these steps to save and close the core file:

1. Hide all sheets ending in "gwl" and the "Main" sheet by right clicking on the sheet name and then selecting hide.
2. Click on the B3 cell in the "Input" sheet and minimize the Excel file (not the Excel program).
3. Click on the Excel save icon.
4. Close Microsoft Excel.

Since the .gwl files still have the same name, no changes are needed in the Evoware script. The Evoware script only prompts the user to select which one of the four worklists he/she wishes to execute and then imports it and executes it.

**For non Tecan Evoware users.** Follow these steps to save and close the core file:

5. Hide all sheets ending in "gwl" and the "Main" sheet by right clicking on the sheet name and then selecting hide.
6. Click on the B3 cell in the "Input" sheet and minimize the Excel file (not the Excel program).
7. Click on the Excel save icon.
8. Close Microsoft Excel.

For non Tecan Evoware software users, there is still one more change to make in order to adapt AG-NGS to a different system. Tecan Evoware uses the worklist extension .gwl, but other systems import files with different extensions. The file extension therefore needs to be changed as appropriate when the worklist is saved. This involves modifying the Excell macro that automatically executes worklist saving and extension assignment. Modify this macro as follows:

1. Double click on the AG-NGS folder and then on the "CoreFiles" folder.
2. Open the TruSeq mRNA sample prep.xls file that has been previously modified.
3. Click on the Programmer menu in Microsoft Excel and then on the Visual Basic button to open the Visual Basic editor.
4. Within "VBAProject TruSeq mRNA sample Prep" select the "Modules" folder and then gwl.
5. Look for lines
   - Open Dir_Forms & "\Worklists\NSG_cDNA.gwl" For Output As 1
   - Open Dir_Forms & "\Worklists\NSG_EndRepair.gwl" For Output As 1
   - Open Dir_Forms & "\Worklists\NSG_LigateAdapter.gwl" For Output As 1
   - Open Dir_Forms & "\Worklists\NSG_Enrich_DNA_Fragments.gwl" For Output As 1
6. Replace gwl with the extension used by your specific software. The name in bold letters will be the name of the file generated by AG-NGS to be imported and executed by the software.
7. Save the VBA script by clicking on the save button in the VBA editor.

Since no two liquid handling software applications are exactly the same, non Tecan Evoware users will need to create a script in their program to prompt the user to select the worklist they wish to import and execute. Check the user manual for your specific software in order to learn how to proceed.

**For all users.** In this case there are four files to be imported, since we have generated four gwl sheets from the "Main" sheet included in the TruSeq mRNA sample Prep Excel file. A file to be imported is generated from each of these four sheets. Generating different files to import allows the user to independently execute different parts of a complex and long protocol (for example running different parts of the protocol on different days).

The next time you open AG-NGS and select "TruSeq sample prep", all the information you enter through the user form will be transferred to the "Input" sheet and will then be used by the user-modified "Main" sheet to generate the corresponding one or more .gwl files (or files with the extension appropriate to your system).

In the event that you need to increase the number of variables entered through the VBA user form, you will have to access the user form and insert new labels and/or text boxes and transfer those values to any empty cells in the "Input" sheet. To access the TruSeq mRNA sample Prep user form, follow these steps:

1. Open Microsoft Excel. Click on "Programmer" and then on the VB button to open the VBA editor.
2. Open AG-NGS to access the main menu. Without closing the main menu, click on the minimized VBA icon in the Windows toolbar to access the VBA editor.
3. Within Userforms double click on "Reagents".

Once you have accessed the user form, click on the VBA help button or check a VBA manual to learn how to insert new text boxes and labels and transfer values to a specific cell in the "Input" sheet. Once any modifications have been introduced, the user form must be saved by clicking on the save icon in the VBA editor.

## How to modify the sample prep for c-Bot script in AG-NGS for cluster generation in the c-Bot.

**For all users.** Sample prep for c-Bot can be modified in exactly the same way as TruSeq sample prep (explained above). The core file for this script is called "TruSeq Dilution samples to seq.xls". In this case there is only one gwl sheet, called "gwl", instead of four.

**For non Tecan Evoware users** the VBA Project for this core file is called "TruSeq Dilution samples to seq", and the line in which the file extension should be changed is

- Open Dir_Forms & "\Worklists\NSG_Denaturated-cBoot.gwl" For Output As 1
- For all users. The user form for this script is called "DilMaker".

## How to modify the Pool Maker script in AG-NGS for pooling samples.

This script is not intended to be modified by the user since it is a stand-alone script for pooling samples regardless of the downstream applications and is not linked to Illumina protocols. However, it is not restricted to Tecan Evoware and can be adapted to any other liquid handling system.

**For non Tecan Evoware users.** In contrast to the other two sub-applications, "Pool Maker" does not generate a "ready to use" worklist but instead generates a .csv file. The .cvs precursor file can be shown in the same way as the "Main" sheet in the other two sub-applications (explained above). The format of this sheet can be modified if necessary to adapt it to the requirements of other systems. Make sure that unused rows are flagged with the word "REMOVE" in column A.

If the file extension has to be changed follow these steps:

1. Open the Pool_Maker.xls core file and then the VBA editor.
2. Within "VBA project Pool_Maker.xls", select modules and then "Guardar_csv".
3. Change csv to the required extension where indicated in red.

```
Sub Crear_csv()
Sheets("Input").Select
Dir_Forms = Range("AA1").Value
Sheets("csv").Visible = True
Sheets("csv").Activate
'Con esta instrucción la celda activa siempre es la E1.
Range("A12").Select
Counter = 96
' Loops through the desired number of rows.
For i = 1 To Counter
    ' Checks to see if the active cell is blank.
    If ActiveCell = "REMOVE" Then
        Selection.EntireRow.Delete
        ' Decrements count each time a row is deleted.
        ' This ensures that the macro will not run past the last row.
        Counter = Counter - 1
    Else
        ' Selects the next cell.
        ActiveCell.Offset(1, 0).Select
    End If
Next i
    If Dir(Dir_Forms & "\Worklists\Pool.csv") <> "" Then
        Kill Dir_Forms & "\Worklists\Pool.csv"
    End If
        ActiveWorkbook.SaveAs Filename:= _
        Dir_Forms & "\Worklists\Pool.csv" _
        , FileFormat:=xlCSV*, CreateBackup:=False
Sheets("Pool").Visible = False
MsgBox "Export Successful", 64
End Sub
```

4. Save the script.

*This last CSV has to be replaced by the appropriate file format, not the extension itself. For example, for .gwl files the file format is "TEXT".

To use the file in software other than Tecan Evoware, the user will need to create a script in the corresponding software to import the file and execute it. Check the user manual for your specific software in order to learn how to proceed.